AD 690507

COORDINATE INDEXING ADAPTED TO OFFICE FILES

Results of Experimental
Application of Coordinate Indexing
to Office Correspondence and Files

A. G. HOSHOVSKY

D D C

JUL

OFFICE OF AEROSPACE RESEARCH
UNITED STATES AIR FORCE

RESULTS OF EXPERIMENTAL APPLICATION

OF

COORDINATE INDEXING TO OFFICE CORRESPONDENCE AND FILES

A. G. Hoshovsky
Office of Aerospace Research
United States Air Force

CONTENTS

NOTICE

This report is published to promote
application of technical knowledge
derived from the Department of the
Air Force research and development
efforts.

The potential users of this report
are cautioned that this document is
not a formal Air Force manual and does
not replace, amend or otherwise
supersede any instructions or directives
issued by the Department of the Air
Force. Those who wish to apply the
principles and procedures outlined
in this report are advised to consult
their records management officials to
determine the appropriateness of their
action.

The views expressed herein are those
of the author, and do not necessarily
reflect the official position of the
Office of Aerospace Research or the
United States Air Force.

ABSTRACT

A two-year experimental application has been carried out to determine the potential of coordinate indexing in retrieval of office correspondence and management documents. The study was carried out in several small staff offices of a military R&D organization.

The results show that coordinate indexing is easy to apply, offers significant operational and retrieval advantages, and appears to contribute to good office morale. These gains are balanced against the increased input costs and the requirement for considerable changes in management and secretarial attitudes.

Coordinate indexing is, therefore, recommended for adoption in all situations where there is a need for rapid and accurate retrieval of documents from multisubject files and where the multiple users require unaided access to those files. Where such conditions do not exist, i.e., where the files contain essentially the same type of documents and where retrieval is infrequent, the increased input costs may outweigh any gained advantages.

I. INTRODUCTION

This report is based on a two-year experimental application of the coordinate indexing techniques to the control of office correspondence and report files. As a document description and retrieval technique, coordinate indexing is based on the idea of concept matching, where sequentially filed documents are indexed by keywords to be later matched (coordinated) with the terms contained in the retrieval question in order to recover needed documents.

Here is how generally coordinate indexing works: The incoming documents are described by keywords, and assigned a unique sequential number. A special record is kept on each used keyword to show which documents have been indexed by this keyword. To locate a document

---

*For a detailed description see Section IV.

the filing clerk determines the keywords of the query, looks for the co-occurrence of the same number on keyword cards, and recovers those documents which correspond to the matched numbers.*

In observing the behavior and practices of file clerks and secretaries, this experimenter was impressed with the reoccurrence of indicators of concept matching processes, albeit without being formally recognized as such an activity. Since the matching occurred only in the secretary's mind (as she compared the words of the request with her recollection of similar words in the recently filed documents) it should be rather clear why this phenomenon has not been adequately recognized and exploited for the improvement of secretarial performance.

This newly gained awareness suggested a hypothesis that the coordinate indexing might also prove to be an easy and natural way for the secretaries to control office files and offer a more effective method for rapid recovery of business correspondence. Along with this we began to suspect that most likely the non-use of coordinate indexing in the office was the result of the clerical profession's unfamiliarity with this technique, as well as the computer-oriented history of coordinate indexing where its principal use was in the area of scientific and technical documents.

In contrast to the above hypothesis there were at that time (1965-1966) certain rather widely held notions that coordinate indexes were all right, but not for office files; that coordinate indexing methods require professional indexers; and that professional help would be needed to start and keep the filing systems going.

To resolve this conflict we decided to carry out an experiment which rested on the following assumptions:

o That the principal difference between scientific and technical literature is in the subject matter of documents rather than in the form of documentation.

o That the average office secretary (GS-4 and 5) should easily comprehend and apply the coordinate indexing.

1

o That file size and the use of
computers would not be the material
considerations for the test of the
success of coordinate indexing as
a method for controlling document
files.

With these assumptions we asked three
secretaries, each a high school graduate
and each reporting to a different super-
visor to assist us in an experiment.
Their files consisted primarily of staff
memoranda, staff studies, copies of
official correspondence and a variety of
technical reports relevant to the staff
functions of the Director of Scientific
and Technical Information in the Office of
Aerospace Research.* They were taught the
basic concepts underlying the process of
coordinate indexing, given examples of the
possible ways to organize their files, and
asked to set up and operate independently
their coordinate indexing filing systems.
If they had questions the experimenter
would answer them but otherwise they were
on their own with complete freedom of
action.

Measurements of the secretaries'
filing and retrieval ability and other
factors were obtained at the beginning, in
the middle, and at the end of the experi-
ment. The findings and conclusions re-
sulting from these tests are shown in
Figures 1 and 2 and described in Section
II.

The preliminary findings were pre-
sented in 1967 in a symposium of the
National Archives** during which one of
the participating secretaries gave a talk
on the methods which could be useful in
replicating her experiences. Because of
a heavy demand for specific instructions,
a special guide was drafted, and the

methods are now being taught in the GSA
training courses on office management.
The pertinent sections of this guide
illustrate how coordinate indexing was
used in the experiment and how it might
work. It is included in Section IV of
this report.

II. FINDINGS AND CONCLUSIONS

The results of the experiment deny
the hypothesis that the coordinate index-
ing methods are not suitable for the con-
trol and organization of nonhomogeneous
and dynamic office correspondence files.
These methods are especially effective in
situations where office functions are non-
repetitive (as in a military headquarters,
or on the staff of a corporation), where
rapid and accurate access to files is an
important requirement, and where there
is a need for unaided access by the office
staff. Since a very large number of the
Air Force staff offices belong to this
category, full or partial application of
coordinate indexing methods holds sig-
nificant promise to improve access to
office files, effect reduction of the
required storage space, and promote the
higher morale that comes from confident
dealing with one's files under pressure
of time and operational demands.

On the other hand, where such con-
ditions do not exist, i.e., where the
files contain the same type of material,
where one is always assured that the file
clerk will be present to recover the
wanted documents, where the retrieval is
infrequent or where there is no urgency
associated with the retrieval task, co-
ordinate indexing should be approached
with great care,*** because the input
costs may outweigh any advantages gained
from the process.

---

*These functions include the responsibility for policy generation, implementation and
surveillance with respect to the command's technical communications programs and
activities.
**National Symposium on Putting Information Retrieval to Work in the Office, Washington,
D. C., May 8-10, 1967, sponsored by the National Archives and Records Service, General
Services Administration.
***This is not to say that coordinate indexing is less useful for files of homogeneous
material, i.e., personnel files where the usual access is alphabetical (by name). The
experiments conducted by the U. S. Marine Corps, have shown remarkable success also in
this type of file. The principal use of coordinate indexing in such cases is mostly to
retrieve data (e.g., give me all the people who are 6 feet tall, have aviation training
and not married) rather than documents.

The above conclusions are supported
by comparisons between coordinate indexing
and subject-heading filing applied to
essentially similar files, as shown in
Figure 1 below.

THE COMPARISON OF AVERAGE

TIME TO EXECUTE FILING-RETRIEVING

FUNCTIONS*

Function	Sub-Function	(N=32) Coordinate Indexing (Time Average) Minutes	(N=10) Subject Heading (Time Average) Minutes
Input (filing)	Decide on index words	3	-
	Decide on subject category	-	1/2
	Record on cards/cross-reference	1 1/2	1/2
	Mark the documents	1 1/2	1/4
	Put away (File)	1/2	3/4
TOTAL		6 1/2 Minutes	2 Minutes
Output (retrieval)	Decide on retrieval words	1	-
	Decide on retrieval subject category	-	1
	Find co-occurrence of words	1/2	-
	Locate the "relevant" files	1/4	13
	Make the final selection	1/4	1/4
TOTAL		2 Minutes	14 1/2 Minutes
Maintenance (Re-filing)	Replace the document	1/4	1
	TOTAL	8 3/4 Minutes	17 1/2 Minutes

*In actual size and composition the files were not identical because files changed
with passage of time.

FIGURE 1

In general the experiment produced the following findings:

Finding One: Coordinate indexing increases filing efforts [input costs].

The average filing time per document shows significant increase over that required in standard subject-heading files; from an average of two [2] minutes to about seven [7] minutes per filed document. This increase is caused principally by the time required to read or scan the document [as opposed to minimal or no reading time in subject heading files] and the general document preparation [assignment of index terms, occasional look up of the thesaurus of index terms and appropriate markings on retrieval cards].

Finding Two: Coordinate indexing drastically reduces retrieval time.

The average retrieval time per document is about two [2] minutes per document. This contrasts with an average of over fourteen [14] minutes for access to the same type* of files organized by subject categories. This is a significant [sevenfold] reduction of waiting time for wanted information, and is especially important where rapid response to queries is required.

Finding Three: Coordinate indexing reduces the rate of retrieval failure.

During the entire period there was only one failure to retrieve the wanted documentation. The total number of requests during this period is estimated at over 500. The subject-heading tests, based on ten [10] retrieval questions, suggest a failure rate of two in eight requests. If these figures are representative, coordinate indexing offers the potential for the reduction of the retrieval failure rate from 25 per cent to two tenths of one per cent.

Finding Four: Coordinate indexing reduces training and cross-training requirements, and facilitates unaided staff use of files.

Coordinate indexing requires less time for training or cross-training of new secretaries than the subject category files. The tests performed on summer employees [predominantly high school students] showed that the training for retrieval takes less than one [1] hour of explanation and two to three practice trials. Training to index** from a thesaurus takes an additional six [6] to eight [8] hours. No comparable tests were carried out for subject heading files; a precise comparison between these two methods cannot be drawn. Two participating secretaries estimated, however, that one needs several weeks of on-the-job training to acquire comparable confidence in dealing with the subject-heading files.

The use of files depends primarily upon one's ability of constructing a search strategy from a thesaurus, rather than knowing the organization and contents of a file. Because of this the coordinate indexing files are particularly well suited for those occasions when the secretary is absent and the staff or the substitute secretary has to look up the filed documents. In our experiment the unaided use of files by the staff when the secretary was busy with other work, was observed on the average of twice per day.

Finding Five: Coordinate indexing adds to the incentives to keep the files small.

*Not much claim is made for the situations where the files reflect the repetitive and precisely defined functions of a given office because the experiment did not include such situations.

**The initial "intuitive" keywords are assigned by the office supervisor, the trainee was only translating these keywords into the appropriate thesaurus code words. A completely independent indexing would take much longer, however, no tests were carried out to determine such training period. An average thesaurus consisted of approximately 150 words.

In our experiment the original files were initially reduced by about two thirds of their original size. In the two years of "experimental" conditions the files were also subjected to frequent purging* which kept their growth to a minimum. Although accurate figures were not obtained, the secretaries estimated that less than 50% of the incoming materials were permanently filed; the majority of "non-consequential" material was kept in a temporary 60-day file or promptly discarded. This freely-elected discipline is explained by the reluctance on part of the secretaries [as well as their supervisors] to devote 7-10 minutes of indexing time to each of the incoming documents. Thus, it appears that a filing system which uses coordinate indexing has a powerful built-in incentive to make immediate decisions about the future value of the documents, and to discard [or temporarily hold] those which are of transitory nature and marginal value. Such incentives do not appear to exist in the subject-heading files where there was no immediate penalty for storing worthless memos and "thank you" letters.**

Finding Six: Coordinate indexing promotes filing confidence and contributes to a sense of well-being.

The reaction of participating secretaries to the coordinate indexing system was found to fall into two categories: 1] a clear preference for co-ordinate indexing approach and 2] studied indifference. The two secretaries who clearly preferred coordinate indexing were young [18-20 years old]*** and employed less than 4 years. One secretary who found coordinate indexing of less help is a mature woman [over 40], with more than 15 years in the Federal employment, and generally acknowledged as an experienced office manager. The two who solidly preferred the coordinate indexing pointed out that the main reason for their satisfaction was the assurance that no matter how imprecise the request, any document in the files could be found. This gave them confidence that they could not fail. Moreover, they felt that the coordinate indexing method allowed them greater participation in the work of their supervisors because now they had to read the correspondence before it was filed away.

Finding Seven: Coordinate indexing reduces overall filing and retrieving efforts.

As can be readily seen from Figure 1, the total time devoted to filing and retrieving each document according to a coordinate indexing scheme is consider-ably less than by the subject-heading system. Our comparisons disclosed that the average time for coordinate indexing was slightly over 8 minutes, while more than 17 minutes was needed for the same operation by the former subject-heading files. If one considers that in an actual operation the usual ratio is five retrievals to each original entry, the saved time becomes more significant and makes the coordinate indexing system even more attractive.

III. THE EXPERIMENTAL METHODS

The Environment

The experiment was conducted in the staff offices of the Director of Scien-

*Coordinate indexing alone should not be credited for this action. Rather, in one case it was the policy that the entire office files should never exceed one filing cabinet; in another, the small office space and periodic decision to "get rid of the junk" were the principle causes of purging.

**This may explain why in government and industry there is a need for annual file reduction exercises, and a legion of records-disposition supervisors to enforce the filing discipline.

***Not much should be made out of the age factor, except as it may be coupled to the general experience and file confidence.

tific and Technical Information, Head-
quarters, Office of Aerospace Research,
consisting of 16 employees in four
divisions. The involved staff [six
professionals and three clerks] responds
to various requests for information about
command policies and practices in
technical communications, develops plans
for new policies and programs, surveys
the degree of compliance with the
command's technical communication and
information transfer policies, and co-
operates with other DOD and Federal
technical information activities in
developing coordinated national techni-
cal information activities. To support
these functions the staff keeps files of
staff studies, policies, plans, and
relevant correspondence. Before the
experiment these files were housed in
seventeen [17] filing cabinets. The
present number is eight [8].*

The Experimental Hypothesis

The experiment was built to confirm
or deny a hypothesis that the coordinate
indexing methods are not suitable for
the control and organization of non-
homogeneous and dynamic office corre-
spondence files. The hypothesis would
be denied if within one year the clerks
succeeded in establishing and satis-
factorily operating their principal
office files using the principles of
coordinate indexing.

The Experimental Conditions

Three secretaries [ages 18, 19,
and 43] with varied backgrounds and
temperaments, all with previous office
experience of not less than three years
were asked to convert their standard
subject-heading files to a filing arrange-
ment based on the principles of coordinate
indexing. The offered choices were
arrangements based on the uniterm cards,
Termatrex cards, and Port-a-punch cards
[See Section IV]. The computer support
of such files was not offered, thus
restricting the girls' choices to manual-

systems. Training in the application of
coordinate indexing methods consisted of
several sessions, each of 30-40 minutes
in duration on a person-to-person basis
between the secretary and the experi-
menter. The girls were formerly exposed
to the coordinate indexing concept be-
cause the experimenter was already using
this method to control his professional
papers. No written guidelines were
furnished, since none existed.

After several discussions the girls
decided to use uniterm cards [Figure 2]
as the most convenient and economical
method. The specific configuration of
each filing system, the construction of
vocabulary, acquisition of special card
holding cabinets, etc., were left to the
discretion of each secretary.

The Procedures Used by the Secre-
taries

Although each secretary was free to
use or not to use any part of the sug-
gested coordinate indexing procedures,
their overall approach followed a gen-
erally similar pattern. This approach
is best described with the excerpts of
the briefing given by one of the secre-
taries to the National Symposium on
Putting Information Retrieval in the
Office.** Here are the pertinent parts
of this briefing:

"a. File Analyses

We began by making a list of
our files, finding out what
our files contained. We
found they contained a good
number of professional papers
that our employers kept, such
as an important journal article
or technical reports to keep
up with their professional
field. There were staff
studies, assessments, tech-
nical proposals to establish
certain information services,
reports of special committees,

---

*These statements should not be interpreted as implying that coordinate indexing was
the sole reason for the reduction. Other factors mentioned in Section II were of
equal if not of greater importance.

**The talk was given by Mrs. Marla Eades, the secretary in the Information Programs
Division of the Director for Scientific and Technical Information.

general inspection reports, and
plans for the scientific in-
formation program. We had
papers on our budget and money
we program for our projects,
copies of contracts and corres-
pondence on the contracts that
we administer. Also we had
directives, regulations,
letters of instruction, clarifi-
cations of policy, as well as
field comments on the regula-
tions which were prepared in
our offices. We used to file
just about everything -
announcements of meetings,
newsletters that were already
kept in an office right next
to ours, memos for the record
as well as our really import-
ant correspondence.

"I was told that I could do any-
thing I wanted, but first I
should try to understand the
basic idea about coordinate
indexing. After understanding
the fundamentals I was to start
trying it out on a small scale.

"I wanted to achieve two things
[1] to have a system that I
could get things in a hurry,
and [2] a system which would
keep me from putting away too
much clutter. I reasoned
that the filing time was not
to be considered as important
as the speed often required
in the retrieval process.

"b. Thesaurus

Before the paper was put into
the file it was indexed with
terms which identified its
contents. As these terms
were listed in a vocabulary
which was called a thesaurus,
but actually was a notebook
which served as a translation
list. We had to construct our
own thesaurus, and in the
beginning it took quite a bit
of work and discussion with
our bosses to learn their
technical language and their

ways of asking for papers.
In the end we managed to get
along with a limited number
of words [150-200 depending
on the size of the file], now
that our vocabulary was estab-
lished.

"c. Filing and Retrieval Aids
     [cards]

Each of the approved keywords
in our vocabulary was re-
presented on a 5 x 7 card, each
with ten columns, numbered from
zero to nine, into which we
would enter the numbers of the
filed documents. The last digit
of the number determined
which column we would enter the
file number. The only purpose
of these columns is to help in
scanning the numbers. All of
the cards were kept in a Kardex
file, with the keywords easily
seen at a glance. The import-
ant parts of our system were:
the filing by number, the list
of approved descriptors, and
the 5 x 7 cards.

"d. Storage

Following are the steps we
used to enter the document
into the file. As a document
was received and the decision
was made to keep it for future
retrieval, it was first given
the next filing number. Then
it was read, usually by the
secretary, who wrote down the
naturally occurring words in
the text best describing the
document. Then she looked up
these words in her thesaurus
and translated them into the
approved keywords. When this
was completed she took the
cards which represented the
selected keywords and entered
the assigned number on each
card. The document was put
in its place in the cabinet,
and the cards back into the
Kardex file. That ended the
filing process.

"e. Retrieval

The retrieval process was just
as simple. Usually the sec-
retary would get a question
which she translated into a
combination of the approved
words. A question like "get
me the letter from HQ USAF
about a program they want to
start next year to train the
old systems analyst at "AFIT",
would be translated into key-
words "USAF", "Program", "AFIT",
and "Analyst". She would then
pull the four cards which re-
presented the keywords out of
the Kardex file and look for
those numbers which would be
identical on all five cards.
Sometimes she would get two
or three documents this way,
but either she or the person
who asked for the document
would quickly recognize which
one was the right one.

"f. Purging

Despite the fact that our new
approach kept a large amount
of material from ever reach-
ing the file, the amount of file
material has increased with
time. To reduce this build-up,
a limit was put on the amount
of material to be kept in the
file.* The files were purged
when the limit was reached.
Since the build-up was slow, the
purge was not required very
often. You had good reason
to eliminate much of the
filing and reduce your files
considerably because it took
more time to file, so we were
more careful about what was
being put into the system.
Under the new procedures a
person reviewing the document
would ask himself, "Will I
have a future need for this?
What will happen if I ever need
this paper and I find that it
has been discarded? Can I
replace it if it turns out
that I have a need for it?"

These questions led to a lot of
material being discarded; a lot
of other material, of interest
for only a short time, was kept
un-indexed in a handy open file
and then discarded when it was
no longer needed.

"g. Conversion From Old Files

For those of you who might be
interested in setting up a
coordinate indexing system
in your own office, I have four
suggestions. In all probability,
you can go the same route that
we, and other offices which have
adopted our system, have gone.
Here are four points which I
think are worthwhile to keep
in mind. [1] Do not attempt
to change your present system
overnight. [2] Start with a
very small portion of your
file, three or four documents
[things that will be asked for
more often than others].
[3] Put in all desired new
material as it is received.
[4] Add material from old
files as it becomes active
[pulled for some reason].
The last point is quite
important. After about a
year, the remnants of your
old file [that which remained
inactive] can probably be
disposed of, and your new file,
of course, will be in full
swing."

The Measurements

A total of 36 measurements were
taken during the two years to quantify
the following activities:

Time to file

Time to retrieve

Failure to retrieve in time**

These measurements were compared to
ten [10] test retrievals, and operations
in the old subject category files,
carried out before the experiment.

─────────────────────────
*Usually the limit was determined by the existing number of filing cabinets in the
office. No new cabinets would be added to accommodate the overflow.

**This time varied according to the urgency of needed response. Usually, when the
document could not be found in two hours the case was considered a failure.

8

The filing and retrieving measurements of the coordinate indexing system were further divided into -

o time to decide on indexing keywords

o time to record on cards

o time to put documents into file

o time to translate the question into retrieval keywords

o time to look up retrieval cards

o time to locate "hot" files

o time to make final selection

o time to refile

Whenever precise measurements were not taken and the experimenter depended on subjective estimates, these are specifically mentioned in the body of this report

IV. COORDINATE INDEXING GUIDELINES

The lessons derived from the experiment were reduced to a handy booklet which could be used in training those secretaries who wish to adopt the coordinate indexing methods. It has been subsequently printed as a technical report and made publicly available under the title Coordinate Indexing: A New Approach to Office Filing, OAR 68-0014. The appropriate portions of this booklet are reproduced here to provide the reader with a better understanding of the experimental environment.

WHY PEOPLE CAN'T FIND RIGHT DOCUMENTS

Many office supervisors complain about their office files when documents cannot be found when they are needed. Most of them blame their secretaries or their filing clerks. Sometimes they suspect that the filing system is not suited for their purpose. Seldom do they admit that they themselves maybe contributing to the problem.

There are many reasons why one's files may not be worth the trouble. Most of them stem from one or more of the following actions:

0 Everything is being filed. The net result of this policy is a filing system bulging with useless junk.

0 A paper is filed under one predesignated subject. The trouble is that practically every paper you file deals with more than one subject.

0 File clerks and secretaries seldom read any of the letters they file. No wonder they look puzzled when you give them your description of the paper wanted.

HOW DOES THE NEW METHOD HELP YOU

0 It forces you to throw away trash before it gets filed. It takes about 5 minutes to prepare a document for filing. With this kind of work you will think twice before filing useless papers. The time you spend on filing is not lost. You will be rewarded when it becomes necessary to get the documents out of files.

0 It permits access to documents from various points of views and descriptions. You combine the right words and presto . . . they lead you to the right document.

0 It forces your secretary to read the document (or at least scan its contents) before it is ready for filing. Your memory will help to recognize what is being wanted.

BASIC PARTS
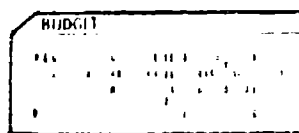
OF THE

COORDINATE FILING SYSTEM

1

STORAGE FOR
DOCUMENTS FILED
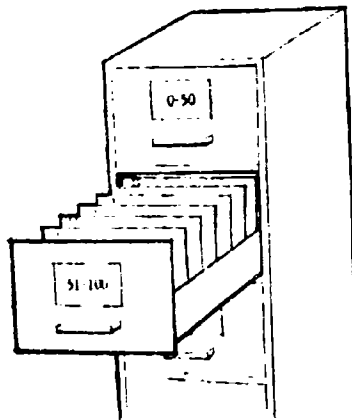
IN NUMERICAL ORDER

2
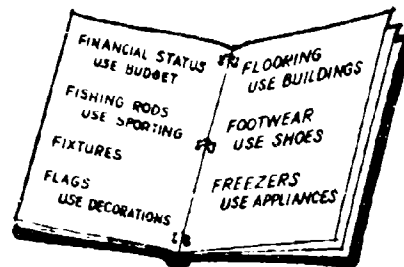
LIST OF INDEXING AND RETRIEVAL
WORDS

3
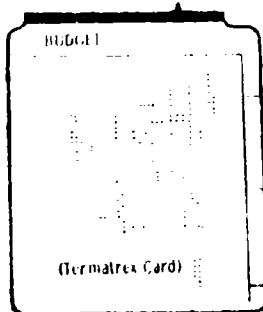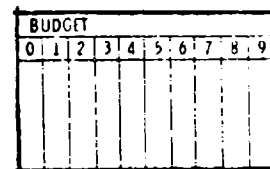
WORD-CARDS
(for recording document numbers)

WHAT IS COORDINATE INDEXING

Coordinate indexing is a new method of filing and retrieving documents with the help of index terms (descriptors.) It has been developed for use in big data storage and retrieval centers where traditional methods of filing by subjects could not keep up with the document volume and the number of requests. A new flexible and faster method was needed. Thus, coordinate indexing was developed to meet this need.

Briefly, coordinate indexing files consist of three parts shown on the facing page.

1. the storage of letters, reports, books, etc., filed numerically, as received for filing.

2. the Word-List is used as a language dictionary. It permits you to start with your own retrieval words and leads you to the *official* file words. (Explained in detail on pp. 14 and 15.]

3. the Word-Card File, containing one Word-Card for each of the *official* words. Record the number of the new document on these cards. [Explained on pp. 16 and 18.]

Here is how the method works.

TO FILE, (1) assign the next number to the new document, (2) choose the Word-Cards which best define this document, (3) enter the document number on each of the selected cards, and (4) file the document numerically.

TO RETRIEVE, (1) decide which word might describe the wanted document, (2) pull the appropriate Word-Cards, (3) look for the same number to appear on all of them, and (4) pull the document from the numbered file.

Since you find your documents by matching, or <u>coordinating</u> the <u>index</u> terms (the descriptor words), the system is called <u>coordinate indexing</u> or <u>concept coordination</u>.
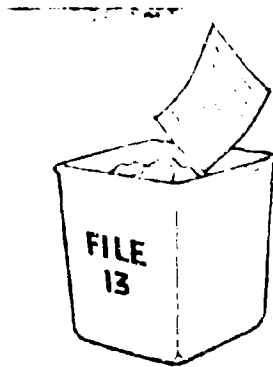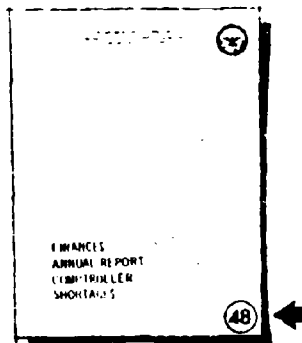
STEP 1

First decide if it is worth filing.
Many memos, thank you letters, etc.,
are not worth keeping for more than
a few weeks, and this can be done in
a temporary folder.

If you decide to keep the document,
also decide how long you will keep it.
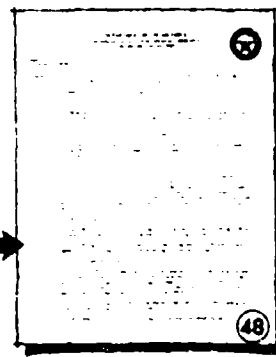(See p. 21.)

STEP 2

Assign the filing number. This
will be the next number in your
files. For example, if your
file has forty-seven project
folders (documents) the next
number will be 48.

STEP 3

Select from your Word-List the words
which best describe the contents of
the document. Pull the corresponding
Word-Cards from your Word-Card File.
Write the words on the filed document.

STEP 4

Mark each Word-Card in the
appropriate place with the
number of the filed document.
(See How To Mark, pp. 16-18.)

Refile your Word-Cards alphabetically.
Put your documents in the file numerically.

12
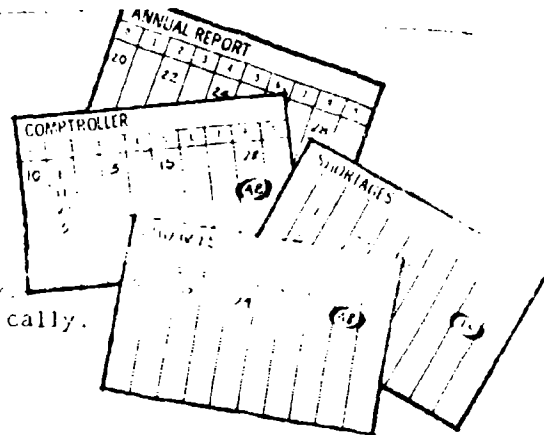
FINDING DOCUMENT IN FILES

STEP 1

Starting with the words of the
request decide which *Word-Cards*
could have been used to index
the document. Use your Word-List
to help you in this decision.

After using the system for awhile
you will find that the need to
use Word-List will slowly dis-
appear. Eventually you will
know most of the words by heart.
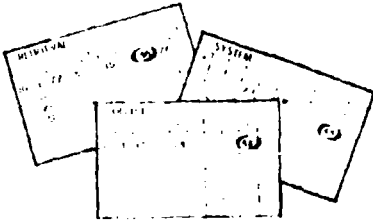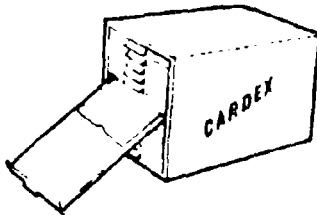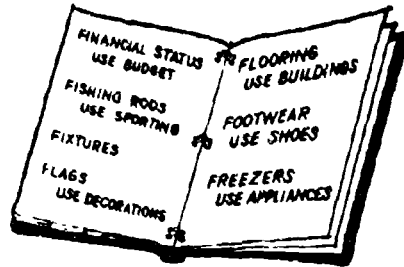
STEP 2

Pull two or three appropriate Word-Cards
from your Word-Card File. Compare the
document numbers. Select numbers which
are identical on all cards.

If you find more than one identical
number on the pulled cards and the
document you want is of recent date,
use the largest identical number. If
it is an older document, start with the
smallest of the identical numbers.

STEP 3

Pull the numbered folder (document)
from files. If more than one folder
has to be retrieved, make your decision
by visual inspection.
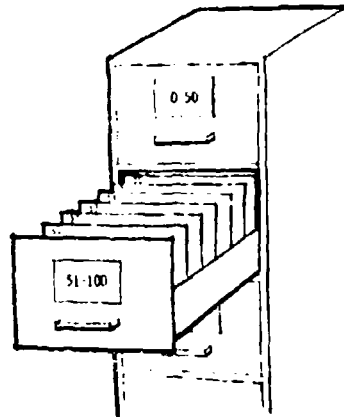
WHAT IS A WORD LIST

The Word-Li: 's a type of translation dictionary. It permits you to
translate con on words used by you or your boss into the *official* filing
words. In a way it acts as a cross reference for words (not files.)
It tells you which words are *understood* by your filing system.

Below is a page from a fictional Word-List used in a retail business.
Notice that its *official* words are capitalized. These are the words
for which there is a corresponding Word-Card in your Word-Card File.
Other words (not capitalized) are those that you might use initially
to ask for the document. They must be first translated into *official*
words. For example, the word financial status might be the word which
your supervisor frequer.ly uses to ask for documents; the document the
supervisor wants is described in your Word-List by the word BUDGET STATUS.
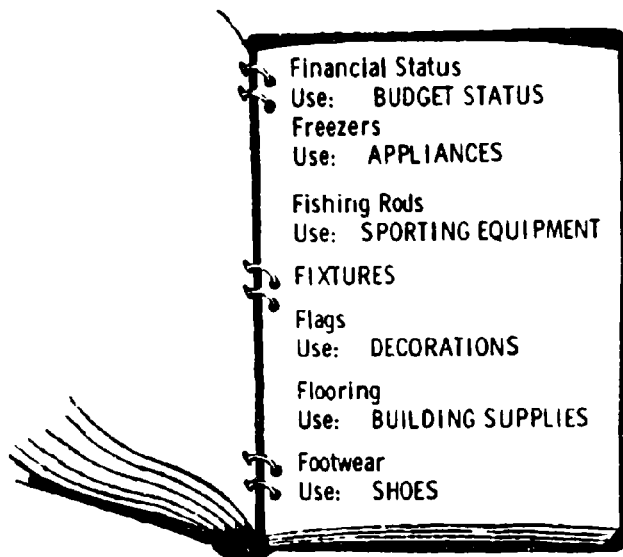
The Word-List should contain most of the words which are actually used
in your office when asking for documents. Included in this list should
be the names of companies and offices with whom you have frequent dealings,
names of certain people, and file control words (shown on p. 21.) The
Word-List can be kept in a loose-leaf binder or on Rolodex cards (see
p. 19.)

HOW TO CONSTRUCT A WORD-LIST

The words you select for your Word-List should be taken from your office files. They should be commonly known and used in your office.

STEP 1

Take about 50 documents from your present files. Circle all words in the title and in the text which describe the document's contents. About 4-7 words per document will be enough.
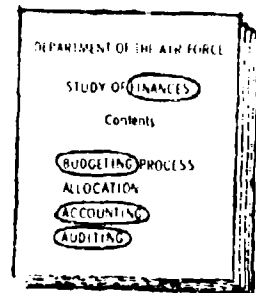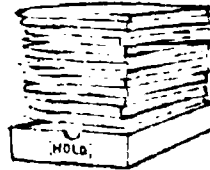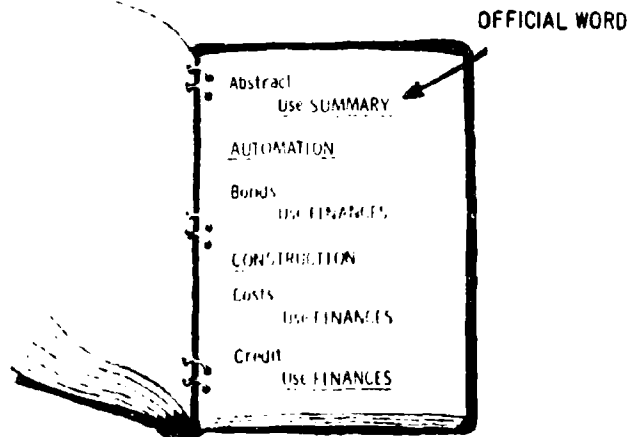
STEP 2

Make an alphabetical list of all selected words. Count how often they have been used. Study the words to see which ones are sufficiently similar to mean about the same thing. Decide which one of these similar words should become your *official* word. The remaining words will be cross-referenced to this *official* word. Too frequently used words or those which are seldom used, make poor official words. Avoid using them as your official words.

WORD-COUNT

Abstract	100
Automation	4
Bonds	13
Broker	2
Construction	32
Costs	21
Credit	9
Debit	1

used too often

insufficiently used

OFFICIAL WORD

STEP 3

The alphabetized Word-List should include both *official* and other words. Type the *official* words in capital letters.

Instead of a typed list you may want to use a Rolodex shown on page 12. This will permit changes in your Word-List without the need to retype the whole list.

Abstract
　Use SUMMARY

AUTOMATION

Bonds
　Use FINANCES

CONSTRUCTION

Costs
　Use FINANCES

Credit
　Use FINANCES

Keep the number of *official* words to a minimum. Remember that for each *official* word you will need a Word-Card.

15

CHOICE OF WORD-CARDS AND THEIR USE

There are three principal non-computerized methods of keeping records
of document numbers in your files: overprinted 5x8 index cards, pre-
scored IBM cards and Jonker Termatrex cards.

The Word-Cards are the principal tools for determining which numbers
in files represent wanted documents.

The numbers are being *coordinated,* i.e., identical numbers are matched
by inspecting the numbers which have been recorded on Word-Cards.

OVERPRINTED 5x8 INDEX CARDS

You can make your own cards by overprinting standard 5x8 index cards.
They are the least expensive and are recommended for the early stages
of your experimentation. If you ever want to move up to the more
sophisticated equipment it won't take much effort to transcribe them
onto either IBM or Termatrex cards.

The numbered columns preprinted on such cards help you to organize
your numbers for easy visual scanning. Notice that the last digit of
the number on the document determines the column where it will be
entered.

By overprinting cards of several colors you can use different colors
for easier access to your Word-Card File. For example, you may want
to use yellow for all action words; blue for all nouns, green for
names of people and institutions, orange for the year or month of
filing, pink for frequently used acronyms, etc. In our tests we
found colors to be of considerable help in pulling the right Word-Cards.

IBM PORT-A-PUNCH C. ᵔ

Special kits of these cards can
be obtained from *Find It*, P. O.
Box 25942, Los Angeles, California
90025 ($10.00 per kit.) The cards
are also available from the IBM
Corporation.
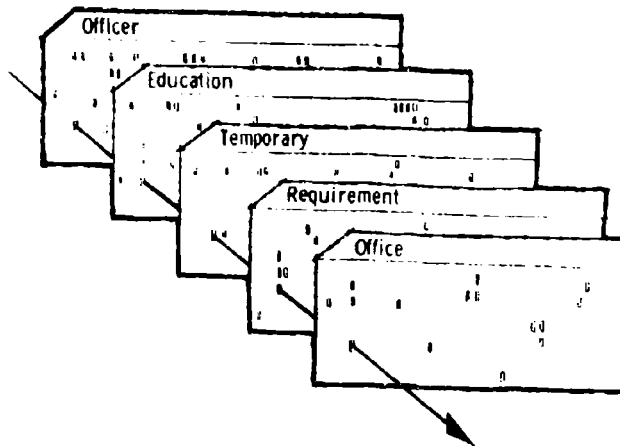
The Port-a-Punch card is manu-
factured by the IBM Corporation.
It is a regular IBM punch-card
on which the holes have been pre-
scored. You can easily push them
out with a pencil or a special
stylus.

The card contains 480 potential
holes, and is suitable for 480
documents. It is most popular
with students and textbook writers
who need to arrange their back-
ground literature for citations,
review, etc.

The card can be *read* by lining
up several cards and letting the
light shine through. It can be
also handled by the standard
electronic data processing equip-
ment.

PUNCH
BOARD
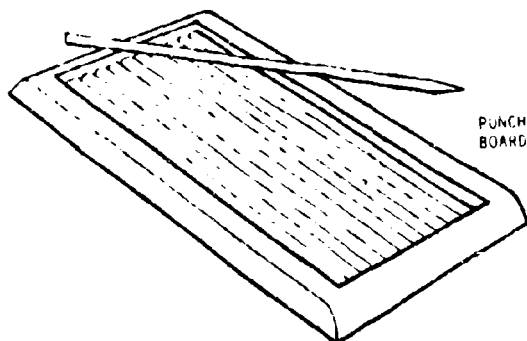
JONKER TERMATREX CARDS

In principle they are similar to
IBM cards, i.e., they are *read*
by passing light through their
holes.

They differ from IBM cards by
their size, material and the
method of making holes. Each
card is 10x10 and contains space
for 10,000 holes. Thus, this
card is suitable for 10,000
documents.

Finding a document number on
Termatrex Cards requires the aid
of a special reading device which
produces a flat light. A special
slide rule permits you to *read*
out the numbers.

LIGHT
BOARD

0048
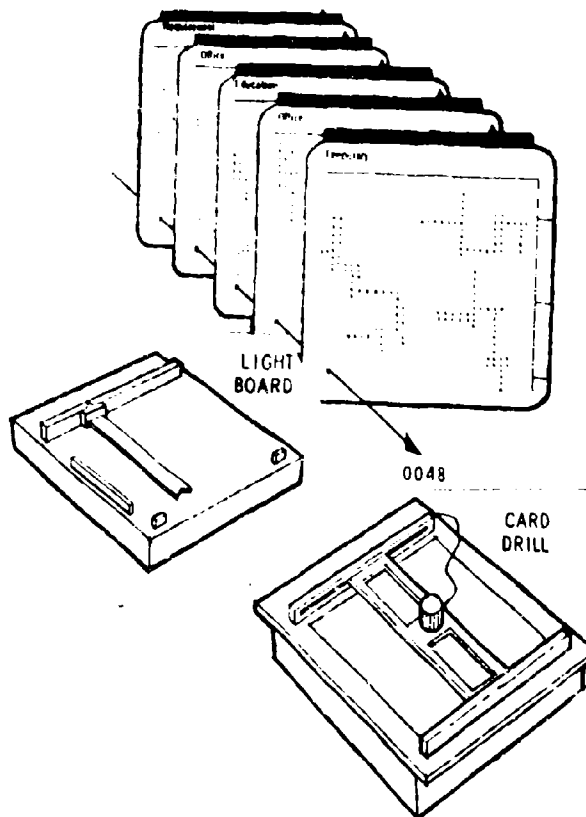
CARD
DRILL

The numbers are recorded by drilling the cards. A special drill is
required to insure that the holes are drilled in precise locations.

The Termatrex System costs more than the other two. Yet, in comparison
with the automatic data processing the Termatrex equipment is quite
cheap. When this report was written, an investment of about $1,500
would buy the basic equipment: cards, drill, light box and card holder.

The systems can be converted to a direct computer read-out with the
help of special Termatrex conversion equipment. The contractual
read-out services may be obtained from Jonker Corporation which
manufactures the equipment.

18

HOW TO UPDATE YOUR WORD LIST

As your files get older you will find that some *official* words are no longer useful, while others should be added. This happens as subjects and interests change. If you do not keep up with these changes your system no matter how good at the start, will become inefficient.

The best way to keep up with the changes is to keep a *candidate word list*. You can record new words as you notice them. After awhile you should review this list, and if you feel that the new words should be added to your Word-List add them.

Remember your Word-List is like a foreign language dictionary. It helps you to translate words used by your boss as he describes the documents. This is why the Word-List should be always accurate to reflect the changes in that language.

Our secretaries found that the standard Rolodex file is the preferred way to keep their Word-List. You can remove or add one word at a time without the need to retype the whole list.

WORD-LIST CANDIDATES
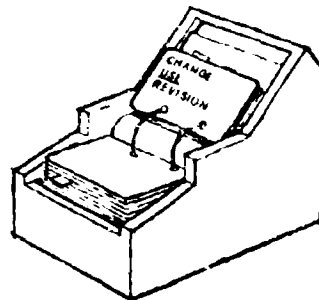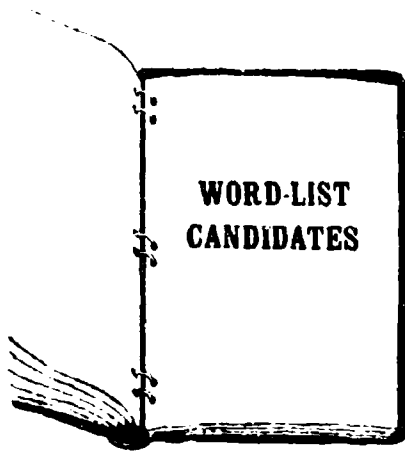
HOW TO CONVERT YOUR OLD FILES

To switch from one filing method to another is not an easy task.
The trouble is what to do with the old files. Re-doing them is
usually a distasteful task, because there are so many old documents
that you just don't see how you can find time to do them.

Lack of time is the principal reason why people continue to suffer
with their old, inefficient method.

FORTUNATELY YOU DON'T HAVE TO CONVERT ALL YOUR FILES. Let's face
it. Chances are good that your old files contain no more than 20%
of useful material which should be kept. The rest is either *archival*,
i.e., should be put away into permanent storage in some central
location, or just simply old and useless documentation that should
be destroyed.

Here is a foolproof way of converting your old files.

1. Keep the old files as they are.

2. As you pull the requested docu-
   ments do not return them to the
   old files. Instead treat them
   as if they were new documents.

3. After one or two years your old
   files will be really old -- and
   dead.

4. Inspect your *dead* files and
   select the *archival* documents.
   Discard the rest. Transfer the
   archival documents to your
   organization's archives.

Using this approach you will be able to convert your files painlessly
and efficiently.

RETIRING OR DISPOSING OF YOUR FILES

Most of the active material in your files is usually less than six months old. It makes good sense, therefore, to dispose or retire the material as early as possible.

Here is one way to tackle this problem:

As you file a document decid in advance about its ultimate disposition. For example, you may decide to keep the document on files for one year and then discard it. Or you may want to keep it two years and then *retire*, i.e., put into the archival storage in your organization.

Establish special *control word-cards* shown in the figure on the left.

Enter the numbers of your documents on the appropriate cards corresponding to your decision.

Once a year review the *one-year* and the *last year date* cards. All numbers that appear on both cards represent the documents that should be discarded. Example 66.

Once a year review the *the year before last* card, and the *archives* card. The numbers appearing on both cards are your archival documents. Example 38.

Once a year review *two years* and the *year before last* cards. Discard the documents whose number appear on both cards. Example 53.

21

DOCUMENT CONTROL DATA - R & D

Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified

1. ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
Office of Aerospace Research Information Studies Division Office of Scientific and Technical Information Arlington, Virginia 22209	UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

Coordinate Indexing Adapted to Office Files

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Scientific - Final - Special

5. AUTHOR(S) (First name, middle initial, last name)

Alexander G. Hoshovsky

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
20 May 1969	21	None

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO. c. d.	 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) OAR 69-0013

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
Tech-other	Office of Aerospace Research 1400 Wilson Boulevard Arlington, Virginia 22209

13. ABSTRACT

A two-year experimental application has been carried out to determine the potential of coordinate indexing in retrieval of office correspondence and management documents. The study was carried out in several small staff offices of a military R&D organization.

The results show that coordinate indexing is easy to apply, offers significant operational and retrieval advantages, and appears to contribute to good office morale. These gains are balanced against the increased input costs and the requirement for considerable changes in management and secretarial attitudes.

Coordinate indexing is, therefore, recommended for adoption in all situations where there is a need for rapid and accurate retrieval of documents from multisubject files and where the multiple users require unaided access to those files. Where such conditions do not exist, i.e., where the files contain essentially the same type of documents and where retrieval is infrequent, the increased input costs may outweigh any gained advantages.

DD FORM 1473 NOV 65

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Indexing						
Records Management						
Retrieval						
Files						
Coordinate Indexing						
Concept Coordination						
Classification System						
Information Systems						